

LES CORPUS EN LINGUISTIQUE ET EN TRADUCTOLOGIE.
MICHEL BALLARD, CARMEN PINEIRA-TRESMONTANT,

Actes des Journées d'Étude « Les corpus en linguistique et en traductologie » organisées par le Certa et le Ceraci à l'Université d'Artois, les 28 octobre 2005 et 27 janvier 2006, Arras, Artois Presses Université, 2007, 346 p.

Alina PELEA

Universitatea « Babeş-Bolyai », Cluj-Napoca, Roumanie
alina_pelea@yahoo.com

S'interroger dans une perspective nouvelle sur des concepts apparemment entérinés dont l'utilité ne fait plus de doute aux yeux des spécialistes s'avère toujours une démarche enrichissante et parsemée de surprises : des questions qui trouvent des réponses, des hypothèses qui se voient confirmer ou infirmer et, surtout, de nouvelles pistes qui s'ouvrent et demandent d'être étudiées. Et si, en outre, l'interrogation vise deux disciplines qui se partagent les notions sans toujours se partager les concepts cachés derrière et les méthodes de travail, le débat devient passionnant.

Tel est le cas du défi qu'ont relevé les scientifiques réunis en 2005 et 2006 autour de la problématique (lire définition, catégories, relevance, limites) du *corpus*. Si tous, linguistes et traductologues, en font usage depuis longtemps, les allers-retours parfois inévitables entre deux disciplines qui restent apparentées en dépit de leur spécialisation croissante exigeaient un « état des lieux » et une mise au point. Car les définitions, les critères de constitution et d'évaluation aussi bien que les usages divergent en fonction des sous-disciplines, des buts, du type de raisonnement. Ces conditions rendent difficile et complexe la réponse à la question-clé des deux journées d'étude, à savoir si la notion de corpus peut être rattachée à la méthode d'analyse utilisée.

Le volume s'ouvre sur un exemple qui ne relève d'aucune des disciplines directement concernées, mais qui les touche pourtant de près. En explorant la façon de Charles Darwin d'appréhender son corpus, Jean Paul Rosaye (« Corpus et modélisation : l'exemple darwinien ») montre combien l'utilité d'un tel instrument dépend du « savoir-faire » du chercheur qui le sélectionne pour ensuite l'interpréter à l'aune des questions qu'il aura définies en début de parcours.

Ahmed El Kaladi (« Corpus et usage de corpus ») fait le passage vers la linguistique en opposant les *corpus linguists* aux *armchair linguists*, puis procède à une systématisation des types de corpus et de l'évolution de la réflexion autour du corpus, pour enfin s'arrêter sur l'exhaustivité, la sacralité, les normes grammaticales et l'interprétation du corpus dans la perspective d'une théorie de l'énonciation.

À partir d'un exemple concret (le discours syndical), Maurice Tournier (« Corpus de textes en lexicométrie sociopolitique ») fait le point sur les étapes et les éléments d'une recherche lexicométrique, tout en attirant l'attention sur les écueils et les risques. « Hardiesse dans les expériences, prudence dans les inférences... », c'est ce que l'auteur recommande à la fin d'une étude qui souligne l'importance d'une bonne hypothèse de travail dans l'exploitation d'un corpus.

En constatant ce qu'il appelle une « Effervescence autour des corpus », car « peu nombreux [...] sont ceux [les] méprisent », Damon Mayaffre choisit de s'arrêter sur un sujet délicat : les limites qu'impose le statut donné à cet outil dans le cadre d'une recherche.

Preuve de la flexibilité des instruments précis de la lexicométrie, l'article d'Émilie Née (« Les outils lexicométriques à l'épreuve d'un corpus médiatique ») exemplifie les étapes d'une recherche sur corpus, tout en soulignant que « l'exploration lexicométrique est [...] un geste inaugural avant d'entrer dans le texte ». Les conclusions de l'étude pratique menée ici au sujet des occurrences du mot *insécurité* dans le journal *Le Monde* sont très riches, mais l'auteure met en garde sur les risques des inférences socio-historiques hâtives : c'est que la lexicométrie peut guider, mais à elle seule elle ne peut pas tout révéler.

Teresa Cabré (« Constituir un corpus de textos de especialidad : condiciones y posibilidades ») aborde le thème des corpus spécialisés, un sujet d'une grande importance, surtout pour la didactique de la traduction technique. L'intérêt de ces corpus est évident, puisqu'ils permettent de donner une caractérisation détaillée des textes spécialisés par rapport à ceux non-spécialisés et, par la suite, d'exploiter les données en linguistique comme en traduction.

Déjà consacrés dans la pratique traductive de même que dans la recherche terminologique, les corpus multilingues n'ont pas encore été pleinement mis à profit. Dans son étude sur l'exploitation d'un corpus trilingue, Maria Zimina-Poirot (« Corpus multilingues : exploration textométrique de l'espace intertextuel ») met en évidence les nouvelles potentialités de la textométrie dans un contexte où la cartographie textuelle ne cesse de bénéficier des avancées techniques.

Aucun texte littéraire n'est dépourvu d'intérêt scientifique, à condition de poser les bonnes questions. C'est la démonstration que

nous fait Volker Mecking dans « L'exploitation d'un corpus du français préclassique ». Son analyse lexicale d'un texte de René de Lucinge, auteur du XVI^e siècle, renseigne avec précision sur « une période charnière de la langue française », tout en ouvrant de nouvelles pistes à exploiter par les historiens de la langue.

Les corpus oraux posent des problèmes spécifiques qui, sinon impossibles à résoudre, sont de toute façon délicats et difficiles. En utilisant le Corpus de Référence du Français parlé, Paul Cappeau (« Constituer et exploiter un gros corpus oral : l'exemple du CFRP ») renseigne sur les aspects théoriques et pratiques liés à ce type de corpus, suggère des approches et met en garde contre leurs limites.

Le passage au volet traductologique du volume est représenté par la contribution de Claude Bocquet, « D'un corpus pour la traductologie à un corpus pour une véritable histoire de la traduction ». L'auteur nuance la définition du corpus en soulignant la spécificité que celui-ci doit revêtir en histoire de la traduction parce que « l'histoire [...] contrairement [...] à la sociologie où à la traductologie, ne vise pas à mettre en lumière des règles de répétition nécessaires ». Les règles de l'histoire générale – donc le regard critique des sources – restent valables pour l'histoire de la traduction aussi.

Dans le cadre de sa démarche de « théorisation réaliste » et pour marquer une « rupture » nécessaire à l'autodéfinition de la traductologie, Michel Ballard (« Étude traductologique sur corpus : la relative dans les traductions d'une nouvelle de Joyce ») entreprend de démontrer que le corpus et son étude acquièrent des valences particulières dans l'analyse des traductions. La linguistique ne saurait être absente du processus, mais elle n'y est qu'un élément parmi d'autres, plus spécifiquement liés au travail pas entièrement prévisible et constant du traducteur.

S'il est courant pour les terminologues d'utiliser les corpus parallèles, cette méthode est moins fréquente et, sans doute, moins systématiquement mise à profit, chez les traducteurs littéraires. Et pourtant, ces derniers aussi bien que ceux qui enseignent la traduction auraient tout à gagner à y recourir, comme le montre John D. Gallagher (« Traduction littéraire et études sur corpus ») à partir du cas concret – et combien sensible – des combinaisons syntagmatiques rares. La validation des équivalences par un corpus n'ira jamais jusqu'à remplacer la créativité du traducteur, mais elle peut certes la stimuler et, plus que tout, offrir des garanties contre les éventuels dérapages.

Dans son analyse des articulations logiques d'un original français, de sa traduction française et d'un texte sur un sujet similaire écrit en anglais par l'auteur de cette traduction, Jean Peeters

(« L'hybridité du texte traduit : corpus et inconscient du traducteur ») vient appuyer l'idée de l'existence d'une langue de la traduction. Démonstration par ailleurs du fait que l'élément étranger ne peut jamais être complètement obscurci, même s'il s'insinue très subtilement dans le texte. Le traducteur fait toujours entendre sa voix, ne serait-ce que très discrètement.

Il va sans dire que les études sur corpus n'auraient jamais pris une telle envergure sans l'existence des moyens techniques appropriés. Franck Barbin (« Statistique textuelle et traduction : quelle pertinence ? ») donne un exemple d'utilisation du logiciel *Lexico 3* sur un corpus très délicat à aborder : dix-sept versions d'une légende populaire de Devonshire. Le traducteur s'attelant à la tâche de la traduire gagnera en rigueur scientifique à utiliser le logiciel mentionné pour mieux maîtriser la variété des originaux en présence.

Adeptes d'une traductologie de l'observation, Corinne Wecksteen (« Le corpus en traductologie : un moyen d'observation pour une approche réaliste de la traduction : application à quelques phénomènes connotatifs ») distingue nettement son domaine de la linguistique contrastive par une analyse des contraintes d'écriture et, par la suite, des décalages registraux vus comme possible « marque de la présence du traducteur ».

Les corpus de traductions spécialisées sont présentés par Mathieu Guidère (« Le traitement des corpus de textes traduits : le cas de la terminologie militaire ») comme un outil intéressant et efficace tant pour les terminologues que pour les lexicographes, les traducteurs et les linguistes. Dans cet article, l'auteur souligne l'intérêt qui existe à les introduire dans les cursus de formation, en vue de consolider « une branche descriptive et applicative de la traductologie ».

Continuant un projet qui avait débouché sur l'identification d'une « distancing trend » dans la traduction littéraire, Adriana Șerban (« Investigating the language of subtitles : questions, methods, applications ») se propose de voir si la situation est similaire en cas de sous-titrage. Sa comparaison du traitement des déictiques dans l'original et dans la traduction indique l'existence de la même tendance, ce résultat étant appuyé par une autre étude récente portant sur les sous-titres anglais d'un film grec.

Trop souvent, les études de traductologie prêtent tant d'attention aux textes, que les destinataires passent au second plan. Or, l'article de Teresa Tomaszewicz, « Études sur corpus audio-visuels et littéraires confrontées aux attentes des récepteurs », vient justement mettre en lumière tout le potentiel du regard des traductologues sur le public réel.

L'étude d'Abdelhaï Sadiq, « Traductique du Coran », clôt le volume par un plaidoyer pour une « théorie des lieux du traduire » et pour une *Transtraduction* qui réunirait les subjectivités de plusieurs traducteurs afin de rendre au lecteur un aperçu fiable du texte d'origine. Sont présentés comme arguments trois versions françaises du Coran : celle de Masson – une chrétienne, celle de Boubakeur – un musulman et celle de Chouraqui – un israélite. Derrière chacune d'entre elles, un destin qui dessine une vision sur le texte d'origine et sur ce que sa traduction devrait être.

Le présent volume est sans doute loin de clore le débat sur le concept de corpus, mais, de par les éclaircissements, les interrogations et les modèles qu'il offre, il constitue un moment charnière dans son déroulement. Les contributions mettent en miroir des approches consacrées et en proposent des nouvelles, sans pour autant négliger le fait que le corpus reste un outil éminemment pratique, ancré dans la réalité de la société. Enfin, cet ouvrage montre qu'au-delà de certains traits quasi-généraux les valences du corpus changent en fonction de la perspective qu'adopte le chercheur.